Recording and registration of new scientific names: a simulation of the mechanism proposed (but not adopted) for the International Code of Zoological Nomenclature

Philippe Bouchet

Muséum national d'Histoire naturelle, 55 rue Buffon, 75005 Paris, France (e-mail: bouchet@mnhn.fr)

Abstract. A discussion draft of the International Code of Zoological Nomenclature which was widely circulated in 1995 proposed that availability of new scientific names would in the future require, in addition to other conditions, their 'international notification' (by listing in Zoological Record (ZR)) within five years of their publication. The application of this proposal (later abandoned) has been simulated retrospectively, to test the criticisms and opposing comments which were expressed by the zoological community. Of 2142 molluscan genus-group names (Recent and fossil, but excluding Cephalopoda) that were established in the period 1980-1992, 260 (12.1%) which were explicitly published as new names were not recorded by ZR; 78% of the omitted names related to fossil taxa. The results highlight the differences between a non-critical recording system and a 'registration' mechanism; the latter would need to evaluate whether and when a scientific name met all the conditions of availability set by the Code. An available name would have to be registered with the accurate date of its establishment, since this determines its precedence. If, in addition to omitted new genus-group names, the unrecorded 'validation' of previously unavailable names and names recorded with an erroneous year of publication or a spelling error are considered, the difference between recording and 'registration' involved 357 names (16.7%). This demonstrates the necessity, as well as the magnitude of difficulty, of establishing a functional and comprehensive registration mechanism for new zoological names. The capture of new names by ZR could probably be improved by some mandatory ruling in the Code, but it is questionable whether a registration mechanism with an acceptably low rate of omission/error can be reached simply as a by-product of routine bibliographical indexing work, i.e. the normal goal of ZR. Any registration, as opposed to recording, system would have to be overseen by the International Commission on Zoological Nomenclature. Neither the Commission nor ZR presently have the capacity to register each year some 20,000 scientific names. However, with modern communication technology, funding through international organizations (e.g., UNESCO) and/or conventions (e.g., the Convention on Biological Diversity) should make it possible to set up a workable registration mechanism early in the 21st century.

Keywords. Nomenclature; taxonomy; registration of names; International Commission on Zoological Nomenclature; *Zoological Record*.

Introduction

A discussion draft of the 4th edition of the International Code of Zoological Nomenclature was widely circulated in 1995. One of the new proposals contained in

the draft was the 'international notification' of new names as a requirement for their admissibility (nomenclatural availability). It was put forward with the following wording:

'Article 11. Requirements.- To be available under the general conditions of availability ..., a name and, where relevant, a nomenclatural act must satisfy the following provisions ...:

- (b) International notification of new names mandatory after [1996]. In order to ensure that the establishment of every new name published after [1996] is notified to zoologists internationally and accessible electronically, for a new name to be available the work in which it is published must be scanned for new names by ZR. Zoological Record is approved for the purposes of the Code by the International Union of Biological Sciences as the record of new names in zoology proposed after [1996].
- (i) A new name recorded as such in *Zoological Record* within five years of its date of publication retains its original authorship and date.
- (ii) If a new name has not been recorded as a new name in Zoological Record within five years of its first publication, it is deemed not to be available from that publication.'

Although the draft Code used the expression 'international notification', the proposal was perceived by the zoological community as a registration mechanism and it elicited two kinds of comments. Some zoologists opposed the principle of registration of new names as a breach of 'the freedom of taxonomic thought or action' which is preserved according to the Code's Preamble. Many others did not oppose the registration of new names in principle, but disagreed with the mechanism put forward in the discussion draft. The most thorough review of the proposal was by Crosskey (1995), who objected on both principle and practical grounds. He wrote as follows: 'This notion introduces into animal taxonomy two principles that have not existed previously: secondary responsibility and temporary availability. [The first involves] shifting onto the shoulders of the indexers/recorders for ZR the responsibility for whether new names shall ultimately live or die, [and] it is hard to see how [the second] new concept can contribute to the stability of names and their authorship and dating. ... Are we to abandon an important name on the technicality that it had failed to appear in ZR within the five-year time frame?'. Crosskey also drew attention to several practical difficulties: (a) determining with accuracy the dates of publication of a new name and of its recording in ZR, and thus whether or not the five-year criterion had been met; (b) ambiguities caused by the appearance of various formats of ZR (paper, disk, online) on different dates; (c) the fact that 'no biological database is ever 100% comprehensive' and that 'to expect ZR to unearth every new name in every publication is quite unrealistic'; and (d) the limited accessibility of ZR to systematists working in disadvantaged countries or locations.

Crosskey's objections were repeated or developed by others. For example, Kerzhner & Starobogatov (1995) said they could give examples of 'works in well-known journals which have not been scanned in five years; of available names listed as nomina nuda, or vice versa'. The subject of 'temporary availability' of new names was also discussed by Rosenberg (1995) and by staff of the Natural History Museum, London (Fortey et al., 1996). The latter emphasized the 'problems in ensuring complete coverage of all new names, particularly those appearing in texts

not using the Latin alphabet'. Many other sceptical comments were made on an Internet discussion forum entitled ICZN-4.

Some of these opposing comments were reviewed by Ride (1996), who proposed that registration by ZR should affect the relative precedence, but not the availability, of new names. Thus, of two available names considered to be synonyms, a name recorded within five years by ZR would have precedence over one not recorded; if neither name had been recorded the dates of publication would determine the precedence (as at present). Ride's revised proposal came late in the discussion process and no comments on it were published in the *Bulletin of Zoological Nomenclature*; some of the objections already made still applied to it and it was not pursued.

In view of the many opposing comments, recording by ZR as a requirement for availability was abandoned by the Editorial Committee and the 4th edition of the Code will not contain such a provision. However, many zoologists think that registration of new names has (or will) become a necessity in view of the mounting diversification of publication sources. Bouchet & Rocroi (1992) documented a 20% rate of omission from ZR for molluscan supraspecific new names published in 1960–1965. Their result elicited controversy (Edwards & Thorne, 1993; Bouchet & Rocroi, 1993; Thorne & Edwards, 1995). At the occasion of the discussion of the zoological Code during the ICSEB meeting in Budapest in August 1996, discussions with the Editorial Manager for ZR inspired me to explore further the recording of new names by ZR. More specifically, a simulation was attempted of what would have happened if the suggested Article 11b [see above] of the 1995 discussion draft of the Code had been in force in the last two decades. I should like to stress explicitly that the purpose of the present work is not to review the accuracy or accountability of ZR, but to contribute to the debate on registration of new zoological names.

Methods

1 have simulated application of Article 11b of the discussion draft to a subset of the new scientific names established in 1980–1992, that is genus-group names of Recent and fossil Mollusca (excluding Cephalopoda). Names published after 1992 had not yet been fully captured by ZR or the Rocroi Index when the study was effected (1996–97) and were therefore not considered. The simulation compares the names recorded in two databases:

- (i) names recorded by ZR, based mainly on the holdings of the British Library and the Natural History Museum, London, together with a small number of donated publications. ZR currently lists 6000 titles as active, and to produce the Mollusca Section it reviews each year an average of some 2400 publications and indexes names from about 2000 of these sources. ZR's policy is to record names according to the way in which they are published, i.e. if an author states that a name is new it will be listed as such, but if the name is presented with an existing author and date, ZR would treat the name as having been previously established.
- (ii) names recorded in a database (thereafter called Rocroi Index) compiled with the assistance of Jean-Pierre Rocroi. ZR has been used as a starting point in the compilation, but other sources are also exploited (see Bouchet & Rocroi, 1992) and access to modern Russian and [former] Soviet literature was facilitated by a working visit to academic libraries in St Petersburg. All names have been checked against the original publication and against the criteria of availability set by the present Code.

After correlation of the names indexed in the two databases, differences (omissions, spellings, dates, authors) between ZR and the Rocroi Index were identified, and the publication source was checked again to confirm (or not) the difference. The study did not evaluate the effect of the 5 year-period proposed in the discussion draft; I have considered all names in the ZR database, irrespective of how long after the original publication the name was recorded.

Results

The Rocroi Index has recorded 2142 genus-group molluscan names that were established (i.e., were made available) in works published in 1980–1992. This list was compared by staff of the ZR with their own database and the deviations of ZR vis à vis the Rocroi Index are shown in Table 1. (Errors discovered in the Rocroi Index are not given here as they are irrelevant to this analysis).

Table 1. Differences between Rocroi Index and ZR

Available names omitted:		281
explicitly proposed as new	260	
not explicitly proposed as new	2	
validation of unavailable names	19	
Names recorded with erroneous date:		64
evidence for error internal	30	
evidence for error external	34	
Names recorded with erroneous spelling:		12
Total:		357

Unrecorded names explicitly proposed as new

Of the 2142 names, 260 (12.1%) that were explicitly proposed as new were not recorded by ZR, i.e. on average 20 new molluscan genus-group names were omitted every year. An examination of the omitted names showed that 46% of the sources containing them are non-periodical publications (books, congress proceedings, and so on), and that these contained 64% of the omitted names (Table 2). This finding confirms the common belief that non-periodical publications are less efficiently captured by ZR (and other records) than are periodicals. Obviously, the reason is that many such publications are not widely publicized and/or are difficult to locate.

Another common belief is that omissions mainly relate to 'obscure' sources and publications in languages using non-Latin alphabets. My findings indicate that China and the former USSR together accounted for 54% of the omissions, but that there were more unrecorded names published in the United States (49) than in the USSR (42). When the literature from 'western' countries (North America, western Europe, Australia, New Zealand) is considered together, it was the source of 50% of the total names and 36% of the number of omissions (Table 3).

When the number of omissions per country (or group of countries) is compared with the total number of names published in that country, we find a very uneven distribution. Nearly a quarter of the new names proposed in the Chinese literature

Table 2. Number of publication sources and genus-group names omitted in ZR for the period 1980–1992 (P = periodicals, NP = non-periodical publications)

	Sources			Names				
	P	NP	Total	%	P	NP	Total	%
USSR	12	7	19	21.3	27	15	42	16.2
E. Europe	3	_	3	3.4	4	_	4	1.5
USA	9	2	11	12.4	21	28	49	18.8
W. Europe	8	10	18	20.2	17	20	37	14.2
Australia/NZ/Canada	3	_	3	3.4	7	_	7	2.7
Japan	2	2	4	4.5	3	7	10	3.8
China	3	19	22	24.7	3	95	98	37:7
Other Asia	2	_	2	2.2	5	_	5	1.9
S. America	6	1	7	7.9	6	2	8	3.0
Total	48	41	89	100.0	93	167	260	100.0

Table 3. Rates of omission of new genus-group names

	Total*	Omissions	% omissions
USSR	412	42	10.2
E. Europe	83	4	4.8
USA	360	49	13.6
W. Europe	505	37	7.3
Australia/NZ/Canada	216	7	3.2
Japan	102	10	9.8
China	395	98	24.8
Other Asia	29	5	17.2
S. America	40	8	20.0
Total	2142	260	12.1

^{*}Total number of new molluscan genus-group names published in literature of stated region in 1980–1992

escaped ZR; with 20% of names omitted, South America came second in rates of omission, but only a small number of names were involved. Contrary to expectations, names in the literature published in the former USSR, eastern Europe, and Japan were not particularly under-recorded but 13.6% of the names published in the USA were omitted. My conclusion is that language of publication, even in alphabets using non-Latin characters (such as Chinese, Japanese and Russian), is not *per se* a source of incompleteness in the recording of new scientific names, which are always written in Latin characters and usually carry identifying labels such as 'gen. nov.' or 'sp. nov.'. In the case of China, for instance, the main cause of omission seems to be the structure of the literature, often involving books and series rather than periodicals.

Finally, and significantly, it may be noted that 78% of the unrecorded new molluscan names were proposed for fossils in works dealing mostly or only with geology and paleontology.

Names not explicitly proposed as new

Twenty-one additional names, meeting the criteria of availability, were omitted by ZR. Kerzhner & Starobogatov (1995) commented on the unintentional establishment of new names. They recognized different kinds, such as premature establishment because the papers of an author or his colleagues appear in an unexpected sequence; or establishment by persons unaware that the name had not yet been published; or, in the case of species-group names, 'upgrading' of infrasubspecific (and therefore unavailable) names. New names established accidentally in keys were also discussed by Noyes (1996). Only two of the 21 omitted but not explicitly new names belong to these kinds of unintentionally established names.

The other 19 are names that previously did not meet the criteria of availability (because no type species had been designated, or no description was provided) and became 'accidentally' available when the missing criteria were met. However, the authors who thus made a name available did not declare it to be new, but merely used it with citation of the original author(s) and date of the earlier publication that had not met the criteria of availability. Such names would not come to the attention of ZR as being new names.

In the forthcoming 4th edition of the Code, a new criterion of availability (Article 16.1) will require that a new scientific name should be explicitly indicated to be new. Failure to comply with this criterion will eliminate those rare instances (such as the two cases mentioned above) of premature or unintentional establishment of new names. However, the 'accidental establishment' of previously published, but unavailable, names will presumably continue and the disqualification of such names (because they are not indicated as being 'new') may cause as many problems as it will solve.

Other problems

Recording by ZR also raises issues of dates of publication. Sixty-four names, i.e. 3% of the total, were recorded with a date that differs from the actual date of publication; additional errors may have escaped my attention. Correct year of publication has been determined by internal evidence in the original publication itself (including statement of exact date of publication published in subsequent issues of a journal) or by external evidence, such as annotations by authors on reprints (generally not available to ZR) or library accession stamps. I should stress that I have considered an 'error' of date to be present only when the calendar year is involved; the precise month and day of publication would be important in a registration system but it has not been considered in the present study.

Finally, there are 12 names (0.6%) that are recorded by ZR with an erroneous spelling, thus leading to the impossibility of retrieving them electronically.

Discussion

Representativeness of the case study

The Mollusca Section of ZR contains the third highest average number of new genus-group names each year, and it is open to discussion whether the omission rate found in the present evaluation based on Mollusca (Cephalopoda excluded) is representative of other zoological groups. The fact that malacology is a discipline

where there are several hundred scientists actively publishing on all continents in many different languages (Bouchet, 1997) speaks in favour of representativity. However, the malacological literature includes a rather large proportion of paleontological literature. In this respect it is probably representative of such zoological groups as vertebrates, brachiopods, corals or ostracods (which are still more dominated by names based on fossils), but it is likely not representative of the majority of terrestrial arthropod groups; the latter account for the larger part of the new scientific names being proposed each year but the proportion in the paleontological literature is smaller.

In addition, management at ZR considers that the period studied (1980–1992) is not representative of their current working practices. In the last ten years, and particularly in the last five, ZR has made significant improvements in the coverage of journals and books and in indexing quality-control (J. Thorne, pers. com.); a new system was introduced in 1993. Of the names published in the last 5 years of the survey period (i.e. 1988–92), only 7.6% were omitted from ZR. This may be evidence of improvement, and ZR believes that this should be even more marked in the next 5 years (J. Thorne, pers. com.). Alternatively, this low percentage of omission may indicate that the Rocroi Index has not yet captured the more 'obscure' names published in the last 10 years.

Taking a 10% overall omission rate (compared with the 12.1% in the present molluscan study) as a working figure, and applying it to the ca. 2000 new genus-group names proposed yearly in zoology as a whole, my results suggest that some 200 names/year went unrecorded in the years under discussion.

Recording vs. Registration

Considering that ZR is by far the most complete indexing source, its failure to record as many as 200 genus-group names each year demonstrates the magnitude of difficulty of establishing a comprehensive recording mechanism for new zoological names. Omissions alone would undoubtedly be a source of nomenclatural instability, as this would affect precedence (and hence the selection of valid names) and homonymy. This certainly gives credence to the idea that registration of new scientific names has become a compelling necessity. However, the present study demonstrates the difference between recording, 'international notification' and registration. As noted above, ZR's recording policy is to index names according to the way in which they are published, i.e. if an author states that a name is new it will be indexed as new, with the date of publication indicated in the publication itself. 'International notification', as specified in the abandoned Article 11b of the draft Code, suggested recording by ZR as a condition of availability. In doing so, it could lead to notification of names that possibly would not meet one of the other criteria of availability set by the Code, or it could notify them with a wrong or inexact date of publication. In other words, 'international notification' would not have liberated a taxonomist from checking whether a notified name is nomenclaturally available and what its date of precedence is.

I believe that the difference between facultative ZR recording and mandatory registration (under ICZN auspices) of new names involves two steps:

- (i) Improving the recording itself, a task that ZR is determined to achieve;
- (ii) Evaluating whether names meet the criteria of availability set by the Code before they are registered, a task which it would be the responsibility of the Commission to

oversee (as is prescribed in the forthcoming 4th edition of the Code for [future] Parts of a retrospective 'List of Available Names in Zoology').

Improving the recording of zoological names

For registration of new zoological names to be voluntarily accepted by the scientific community, its mechanism must be perceived to be handling equally fairly all branches, subdisciplines and areas of practice of zoology. In this regard, the now-abandoned mechanism proposed in the 1995 discussion draft of the Code made several mistaken assumptions.

- (a) The proposal assumed that Zoological Record is the universally used bibliographical index and that a Recommendation to send published materials to ZR for international notification would suffice to bring names to the attention of recorders. Whereas ZR is almost certainly more widely used by animal taxonomists than any other bibliographical service, especially by zoologists in developed countries, this is probably not the case with paleontologists, especially in China, the former USSR and economically less favoured countries. Such scientists might perhaps have little incentive to follow a Recommendation of the Code advising authors to draw to the attention of the ZR any new name published. General Recommendation 24 of the current Code already recommends authors to forward copies of their works to ZR at 'the earliest opportunity'; in practice, very few authors send reprints, but those who do come from many different countries (including China and Russia), suggesting that compliance with the Code is a function of individual preference or knowledge and is independent of country of origin.
- (b) The proposal assumed that 'obscurity' and linguistic difficulties are the main reasons why new scientific names escape the nets of ZR. Indeed, most zoologists seem to accept the idea that, considering the explosion of the scientific literature, authors have a responsibility to make their work visible and known to the community at large. In other words, authors who publish their work in really obscure outlets cannot complain if their new scientific names escape recording by ZR. This is probably what Holthuis (1996) had in mind when he expressed the view that 'The objection that the ZR is incomplete is true, but this is mainly the fault of authors'.

The present work demonstrates that several factors combine their effects to explain the omissions and account for 'obscurity'.

- (i) Although paleontological material is regarded by ZR as part of its field, geological material is not at the core of ZR coverage and any new zoological name published in an otherwise purely geological serial (or, worse, book) would be regarded as 'obscure' in these terms. This may explain why, as mentioned above, 78% of the unrecorded molluscan names had been proposed for fossils in pamphlets, books, serials or periodicals dealing mostly or only with geology and paleontology.
- (ii) What may appear 'obscure' to, e.g., a western European zoologist may be mainstream literature to a Chinese paleontologist. Many of the Chinese books containing new names unrecorded by ZR have been published by Academia Sinica or its branches, or government publishing houses, and the new names in them were recorded by *Gushengwuxue Wenzhai* [Paleontological Abstracts], a quarterly published by the Academy's Institute of Geology and Paleontology in Nanjing, China. However, much of this material is hard to obtain without focused bibliographical research. For instance, I spent two weeks in academic libraries in St Petersburg

specially for the purpose of nomenclatural indexing, and I correspond with the library of the Institute of Geology and Paleontology in Nanjing on a regular basis. Almost all of the Chinese books containing 95 (of the total of 260) omitted names were still unavailable to ZR when checked at the end of 1997. Clearly, better access to this type of literature alone would improve ZR coverage markedly.

(iii) Omissions occasionally affect names published in periodicals, and non-periodical serials, which are normally scanned by ZR. In trying to locate the 89 publications containing names omitted from ZR (Table 2), it was found that 54 were present in the libraries used by ZR, and the names in them were therefore truly overlooked. Regrettably, errors occur in any human system and publishing a work in a serial normally scanned by ZR does not guarantee that a new name will be recorded, or that it will be recorded with its proper spelling and date. This defeats the principle of automatic registration advocated by Rosenberg (1995), and supports Crosskey's (1995) criticism of 'shifting onto the shoulders of the indexers/recorders for ZR the responsibility for whether new names shall ultimately live or die'.

Informal discussions with zoologists and paleontologists suggest that, to be acceptable, a recording or registration mechanism should have a rate of omission/error not higher than 5%, possibly as low as 1–3%. This is an ambitious goal but given a little extra help from taxonomists it would be achievable. The extent to which capture of new names by ZR can be improved by voluntary or mandatory ruling in the Code remains speculative. Considering the amount of omissions of names in Chinese and Russian literature, an avenue to be explored would be the formal involvement of bodies such as China's Academia Sinica or Russia's Akademia Nauk in the indexing process.

Evaluation before registration?

Registration, if any, would be the responsibility of the Commission. However, considering the available resources, the magnitude of the task is daunting: if all names regulated by the Code (i.e. from subspecies to superfamilies inclusive) are considered, ca. 20,000 new names are proposed each year. Clearly, considering that ZR already indexes 88% of the new genus-group names, it is obvious that the zoological community and Commission should build on ZR, rather than attempt to start a wholly new 'registration office'. Malicky (1996) proposed a new Recommendation whereby 'editors of journals and books should be responsible for notifying new names in accepted taxonomic manuscripts to the ZR staff, who would immediately allocate a reference number to each name. This number would be published with the name, thereby informing readers that the name had been recorded; if a name had no number every reader would know that it should be brought to the attention of ZR'. This proposal would lead to labour-intensive bureaucracy and contains several undesirable or unpractical aspects, not the least being that such a mechanism would register names a priori rather than a posteriori (as would be appropriate). But I believe it points the way to the future of scientific name registration.

Zoologists may perhaps soon be in a position to benefit from the experience of botanists. It has been proposed that, subject to ratification by the XVI International Botanical Congress (St Louis, 1999), new names of plants and fungi will have to be registered in order to be 'validly published' after I January 2000 (Borgen et al., 1998).

During the current test and trial phase (1998–1999), all new taxa, all new combinations or rank transfers are registered by the International Association for Plant Taxonomy (IAPT) Secretariat either (a) by being published in an accredited journal or serial, or (b) by being submitted for registration either directly or through a national registration office, or (c) (during the non-mandatory trial phase only) as a result of scanning of other published information by the registration centres' own staff. The test and trial phase also addresses issues such as registration date and acknowledgement to the submitting author that registration has been effected.

Neither the Commission nor ZR presently have the capacity to register yearly 20,000 names. However, the now general use of computers, communication via the Internet and possible funding through international organizations (e.g., UNESCO) and/or conventions (e.g., the Convention on Biological Diversity) should together make it possible to set up a workable registration mechanism for zoological names early in the 21st century.

Acknowledgements

This article was written following extensive discussions with Joan Thorne of BIOSIS U.K., Editorial Manager for *Zoological Record*. It would not have been possible without the collaboration of Jean-Pierre Rocroi, who compiled the molluscan database and provided the data for the present evaluation.

References

Borgen, L., Greuter, W., Hawksworth, D.L., Nicolson, D.H. & Zimmer, B. (IAPT Executive Committee). 1998. Proposals to implement mandatory registration of new names. *Taxon*, 47(4): 899–904 (see also *Biology International*, 36: 34–36).

Bouchet, P. 1997. Inventorying the molluscan diversity of the world: what is our rate of progress? *The Veliger*, **40**(1): 1–11.

Bouchet, P. & Rocroi, J.P. 1992. Supraspecific names of molluses: a quantitative review. *Malacologia*, 34(1-2): 75-86.

Bouchet, P. & Rocroi, J.P. 1993. The lottery of bibliographical databases: a reply to Edwards & Thorne. *Malacologia*, **35**(2): 407–410.

Crosskey, R.W. 1995. Bulletin of Zoological Nomenclature, 52: 229-232.

Edwards, M.A. & Thorne, M.J. 1993. Reply to 'Supraspecific names of molluscs: a quantitative review'. *Malacologia*, 35(1): 153–154.

Fortey, R.A. (Chairman) et al. 1996. Bulletin of Zoological Nomenclature, 53: 15-17.

Holthuis, L.B. 1996. Bulletin of Zoological Nomenclature, 53: 83-84.

Kerzhuer, I.M. & Starobogatov, Y.I. 1995. Bulletin of Zoological Nomenclature, 52: 296.

Malicky, H. 1996. Bulletin of Zoological Nomenclature, 53: 11.

Noyes, J. 1996. Bulletin of Zoological Nomenclature, 53: 11-12.

Ride, W.D.L. 1996. Bulletin of Zoological Nomenclature, 53: 6-7.

Rosenberg, G. 1995. Bulletin of Zoological Nomenclature, 52: 300.